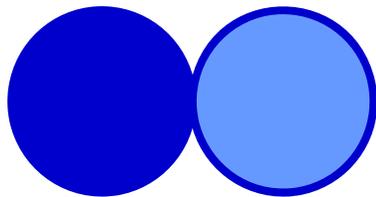
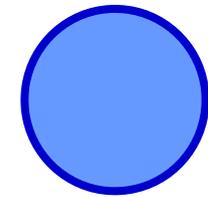
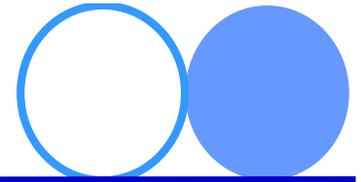
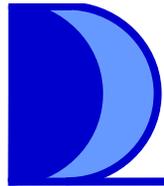


因子分析・確証的因子分析

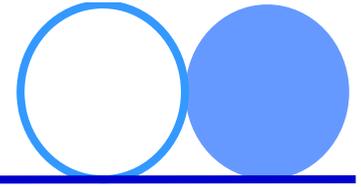
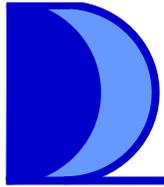


心理データ解析演習 2011/5/18

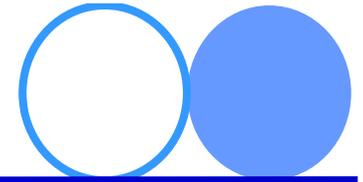
D1 小山内 秀和



- 因子分析とは: 分析を行う前に
- 探索的因子分析: SPSSを使ってみよう
- 確証的因子分析: Amosを使ってみよう



分析を行う前に
因子分析とは



- **因子分析** (factor analysis)

さまざまな観測変数 (= 尺度への回答など) の相関関係から, その背後に共通して存在する, 観測変数に影響を与えているような潜在変数 (= **因子**) を特定するための分析手法

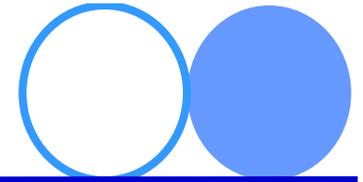
- 「潜在変数→観測変数」という因果関係を推測したうえで, 潜在変数を特定するための手法

- **観測変数**

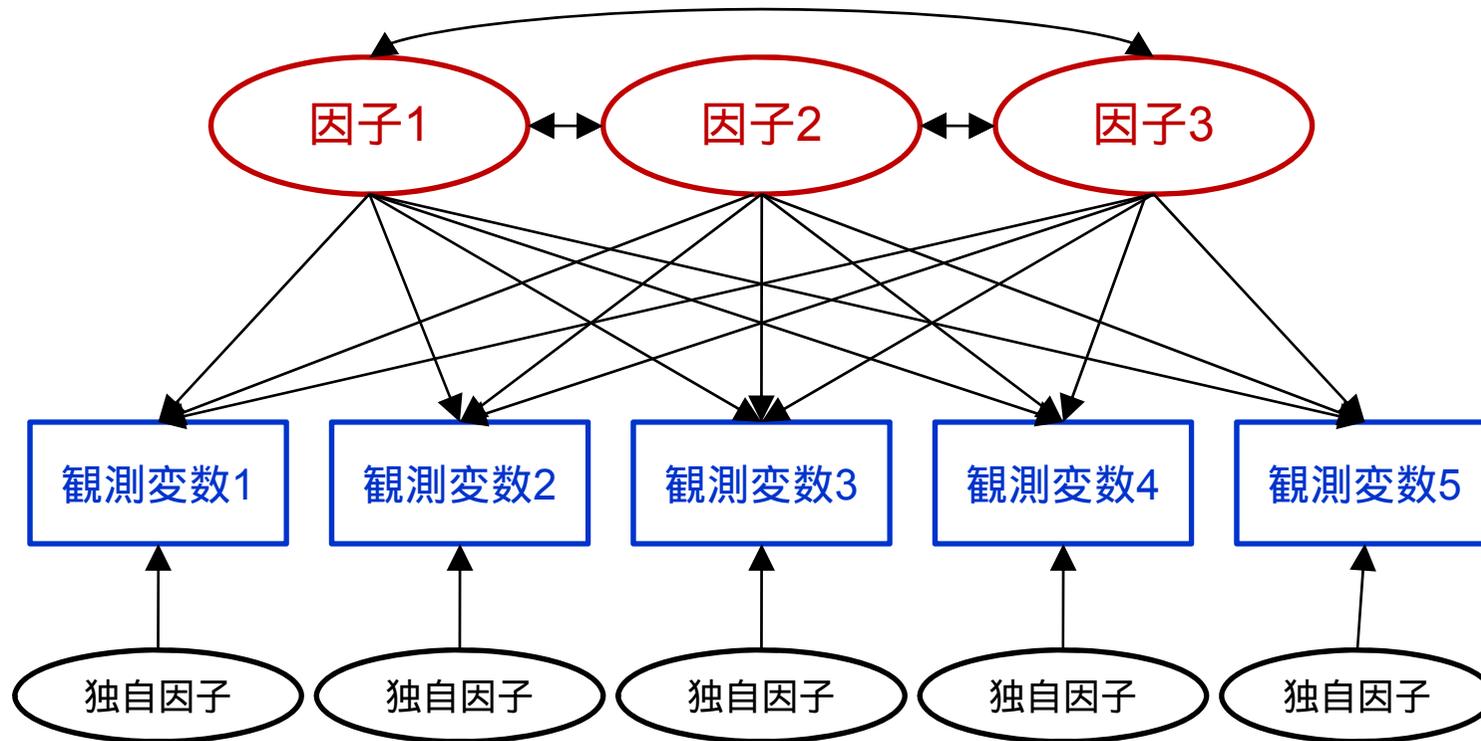
- 調査や実験などによって直接測定した変数

- **潜在変数**

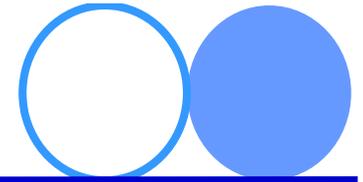
- 直接測定はしていない(できない)が, 観測変数のパターンを説明できるような変数



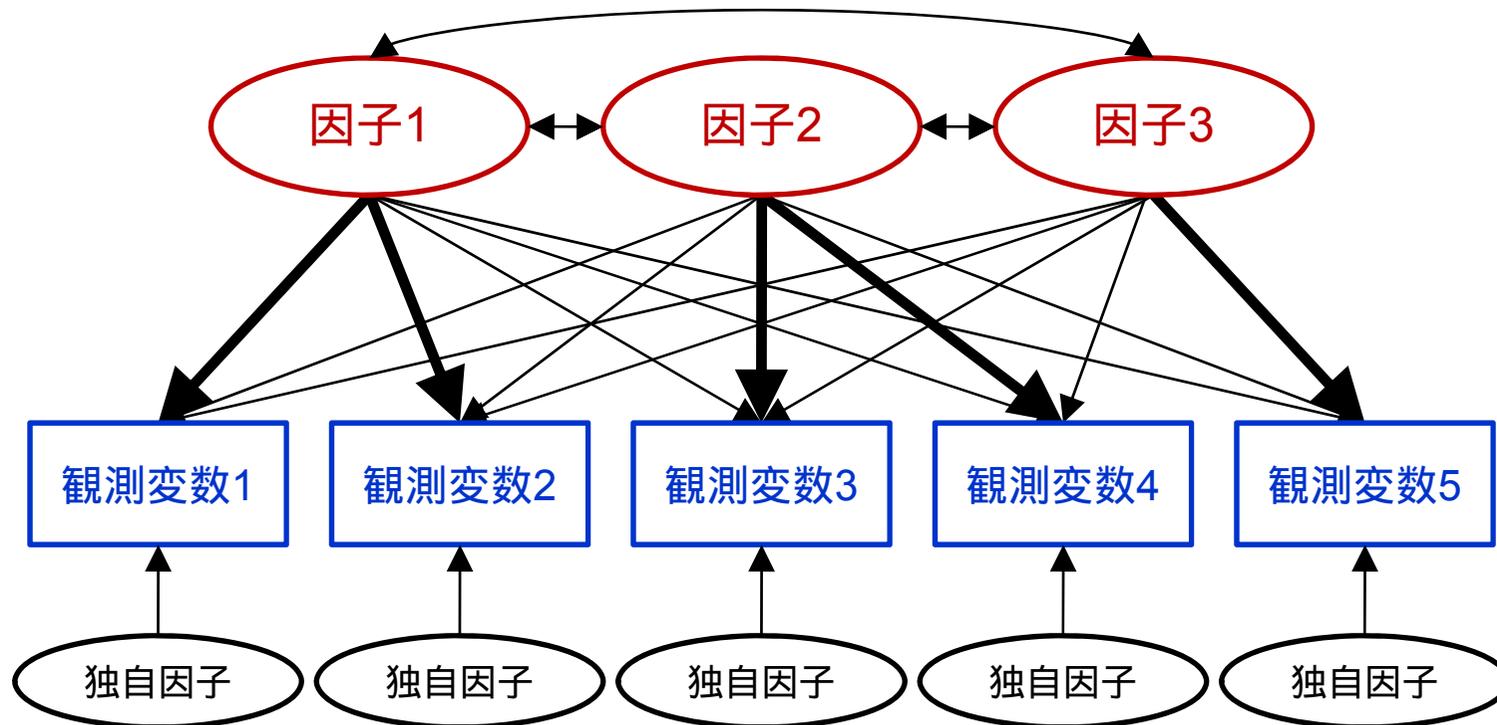
各因子(潜在変数)から観測変数への影響を仮定する

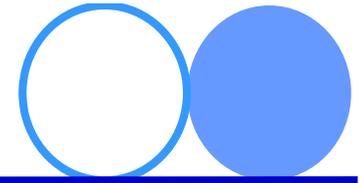


各観測変数は、各因子(共通因子)と独自因子からの影響を受けている



共通因子が、どの観測変数に対して強い影響を持っているか分析する

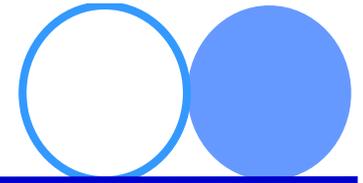
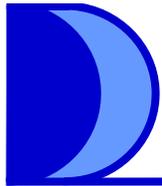




- パス図のうち, 任意の観測変数 (y_i) について式で表すと

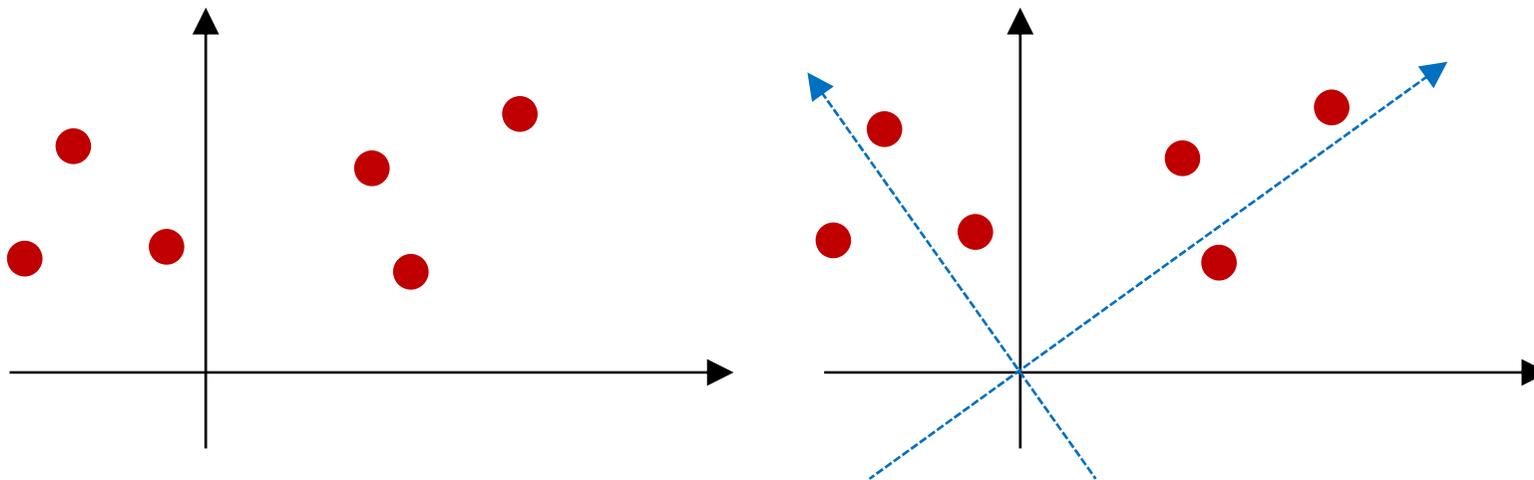
$$y_i = a_{i1}f_1 + a_{i2}f_2 + a_{i3}f_3 + e_i$$

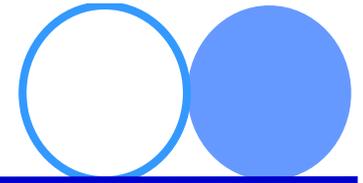
- ただし, $f_1 \sim f_3$ は各 (共通) 因子, $a_{i1} \sim a_{i3}$ は各因子の y_i に対する影響度を表す係数 (因子負荷量), e_i は独自因子
- 各因子について平均 = 0, 分散 = 1 と仮定して計算
- 観測変数 y_i を平均 = 0, 分散 = 1 として計算したものが, SPSS などでは表示される
- 共通性
 - 独自因子 (e_i) の分散を 1 から引いた値
 - 各因子負荷量の平方和に等しい (直交解の場合)
 - 各項目 (y_i) の分散が, 共通因子全体の影響をどれだけ受けているかを示す



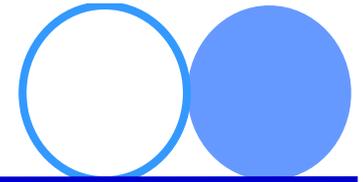
- 「**回転**」とは
 - 「因子→観測変数」式の解を決定する手続き
 - 因子による観測変数の説明率を大きくするための手続き

因子を座標軸として考えてみると…

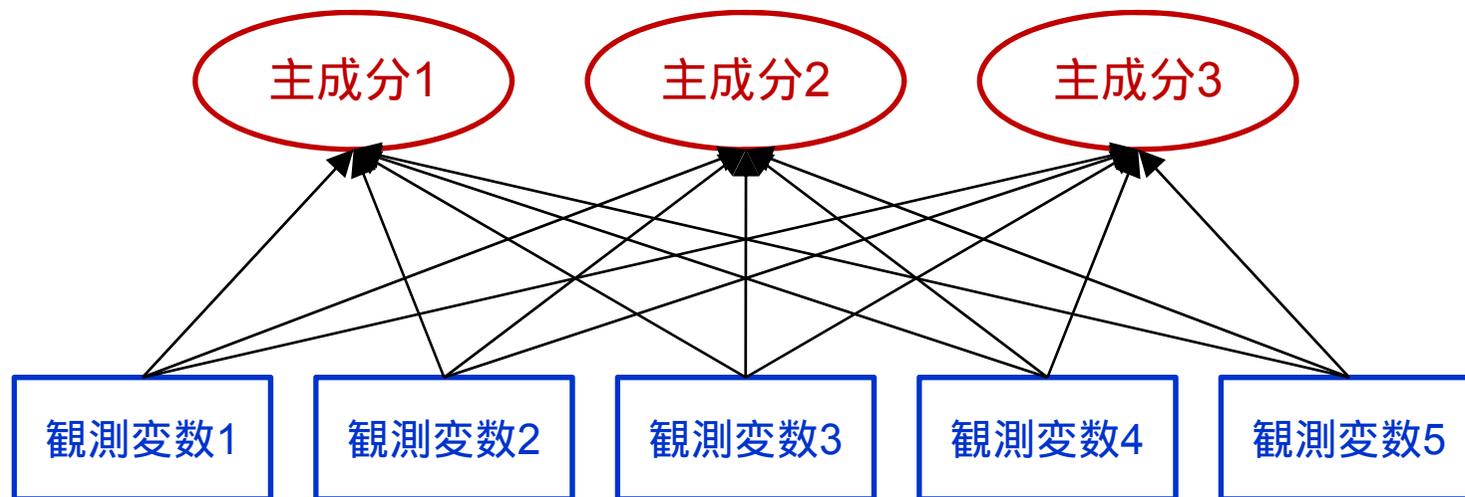




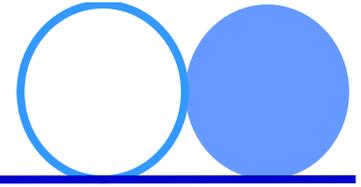
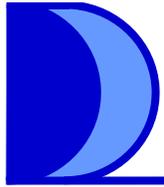
- 直交回転
 - 因子間の相関がない(因子軸の交点が直角)と仮定する回転 → 直交解
- 斜交回転
 - 因子間の相関を仮定する(因子軸の交点を直角に固定しない)回転 → 斜交解
- 現在では斜交回転のほうがよいとされている
 - 以前もそう考えられていたのだが, 斜交回転のほうが計算が複雑だったため, 実質的にできなかった



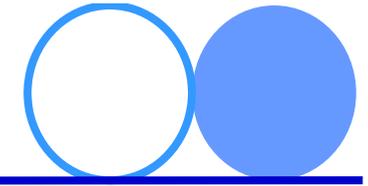
- 主成分分析 (principal component analysis)
多数の観測変数を少数の変数にまとめられるように、
新しい潜在変数 (主成分) を合成する統計手法
- 主成分分析のイメージ



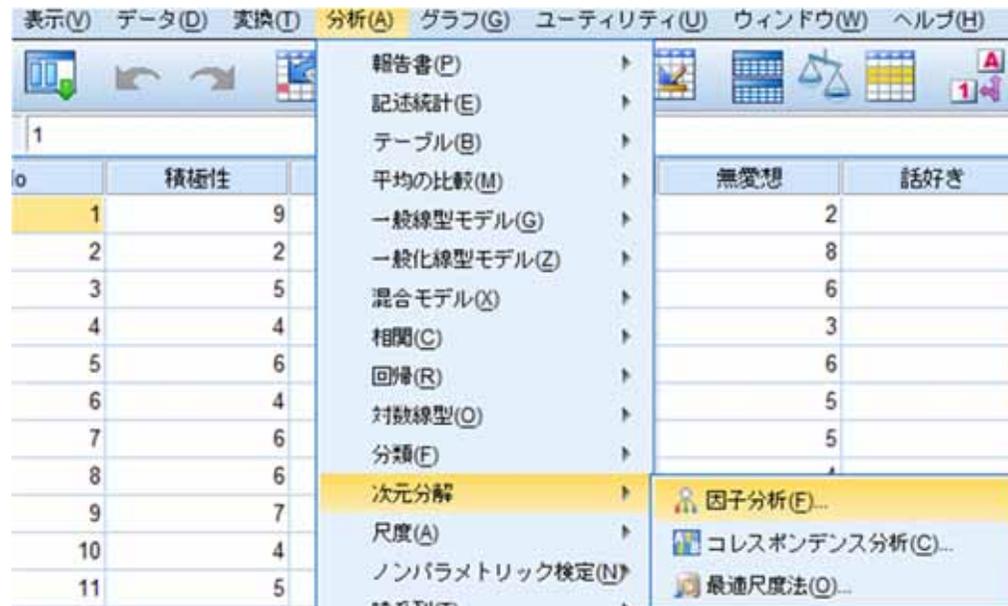
因子分析とは因果のベクトルが逆



SPSSを使ってみよう
探索的因子分析



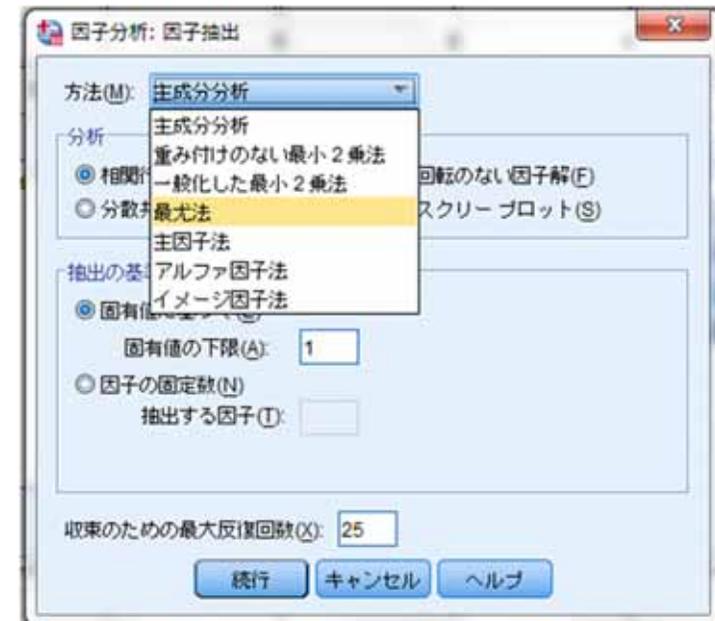
- デモデータ (Excelファイル形式) の読み込み
 - 「性格データ」 (足立, 2006 p. 76 表8.1 (A) より)
 - 「ファイル」 → 「開く」 → 「データ」
- 「分析」 → 「次元分解」 → 「因子分析」

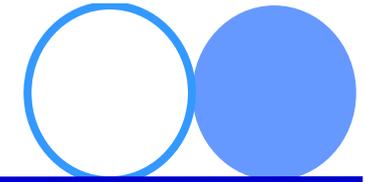


- 分析を行う変数を選択
 - 右ウィンドウへ移す



- 因子抽出法の決定
 - 「因子抽出」をクリック
 - 右画面が表示
 - 「最尤法」を選択
 - 「スクリープロット」にチェック
 - 「続行」





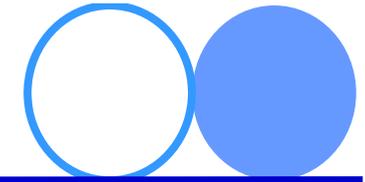
- 回転法の指定

- 「回転」をクリック
- 「プロマックス」を選択
- 「続行」



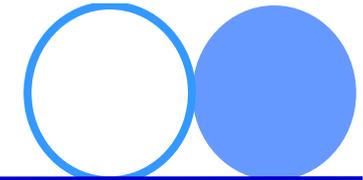
- よく使われる回転法

- 「バリマックス (varimax)」: 直交回転
- 「プロマックス (promax)」: 斜交回転



- 「オプション」をクリック
 - 「サイズによる並び替え」にチェック
 - 「続行」
- 「OK」で分析実行





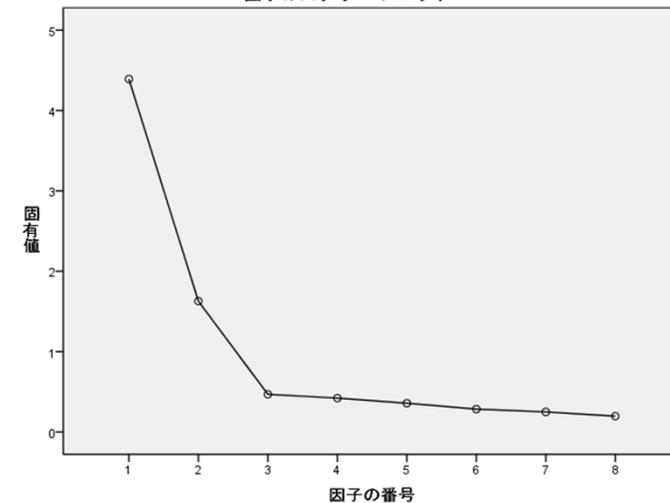
- **共通性の確認**
 - 共通性の値が1以上,あるいは0に近い値でないか注意
- **スクリープロット(回転前)の確認**
 - 因子数を決める参考
 - 固有値1以上の因子数
 - 固有値の折れ線が大きく曲がる部分
- **適合度の確認**
 - 最尤法を選択すると,適合度検定が行われる
 - 「有意確率」が.05以上
 - 「モデルがデータに合っている」という帰無仮説が棄却されない

共通性

	初期	因子抽出後
積極性	.663	.736
陽気	.543	.621
先導	.539	.585
無愛想	.630	.702
話好き	.697	.815
やる気	.680	.777
躊躇	.554	.612
人気	.506	.538

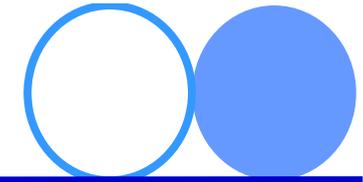
因子抽出法:最尤法

因子のスクリープロット



適合度検定

ガ2乗	自由度	有意確率
8.127	13	.835



- **パターン行列の確認**

- 各因子への回転後の負荷量が表示される
- どの項目が高い負荷量を示したかを見る
- 因子が何を示すかは研究者が判断して決める
 - 第1因子は「活動性」
 - 第2因子は「社交性」

- **因子間相関の確認**

- 斜交回転なので当然確認しておく

因子相関行列

因子	1	2
1	1.000	.504
2	.504	1.000

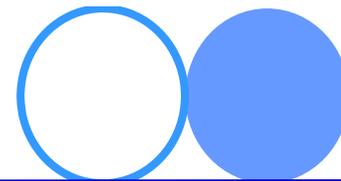
因子抽出法: 最尤法
回転法: Kaiserの正規化を伴うプロマックス法

パターン行列^a

	因子	
	1	2
やる気	.881	.001
躊躇	-.828	.100
積極性	.824	.063
先導	.752	.025
話好き	.021	.892
陽気	-.157	.856
無愛想	-.024	-.825
人気	.224	.595

因子抽出法: 最尤法
回転法: Kaiserの正規化を伴うプロマックス法

a. 3回の反復で回転が収束しました。



- Rを使った因子分析

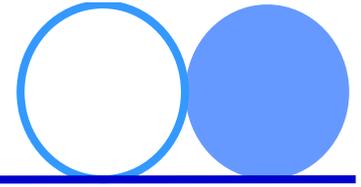
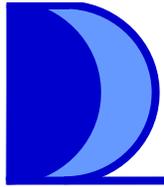
- 「factanal」関数を使う

- ただし, 「rotation = “promax”」と指定しても, このままでは因子間相関が出力されない

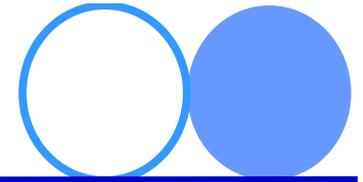
- 因子間相関を算出する方法(青木, 2009)

```
> 代入値1 <- factanal(データフレーム名, factors =  
因子数, rotation = “none” )  
> 代入値2 <- promax(代入値1$loadings)  
> 代入値3 <- 代入値2$rotmat  
> Solve(t(代入値3) %*% 代入値3)
```

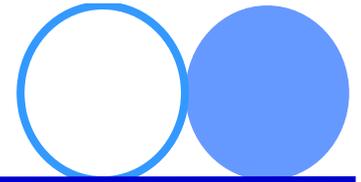
- 「factanal」を使ったときと(基本は)おなじ計算結果で, なおかつ因子間相関を出力してくれる



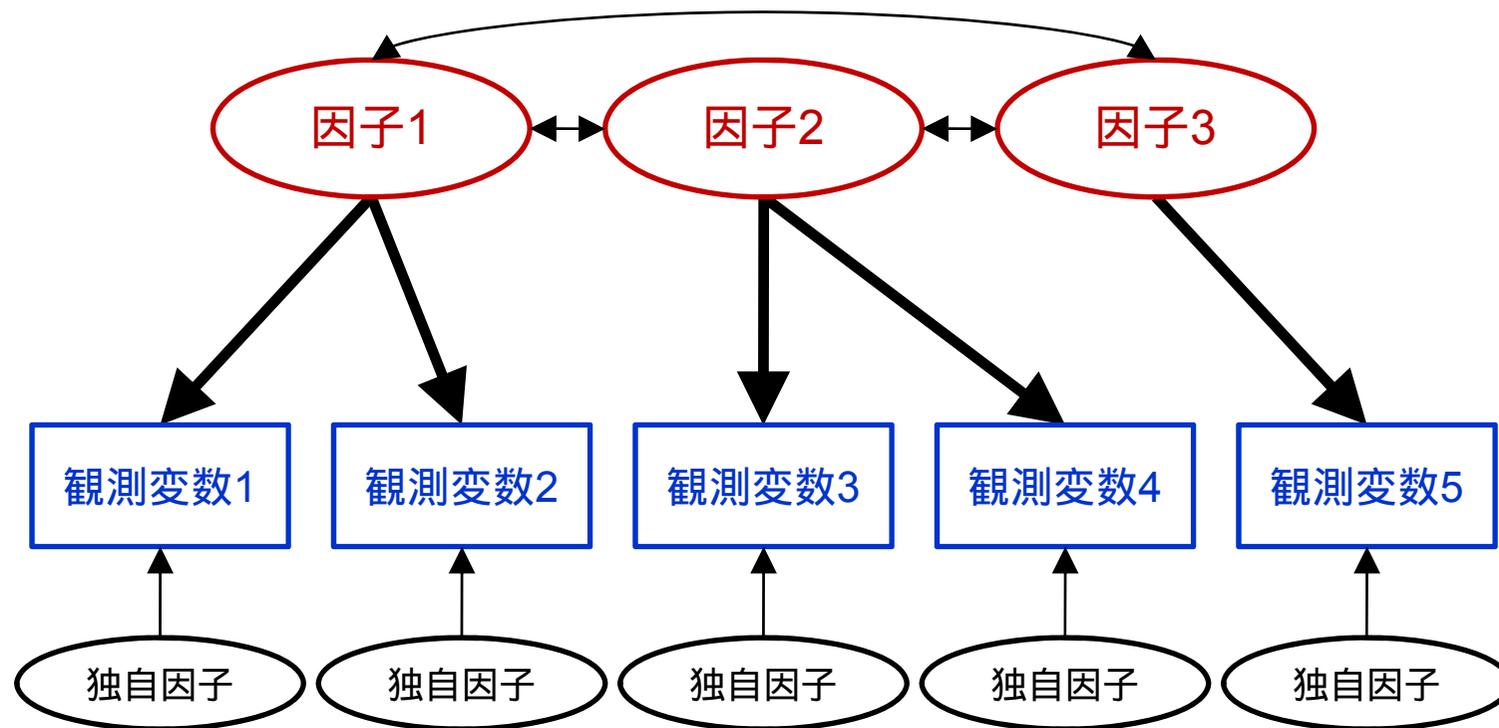
Amosを試してみよう
確証的因子分析



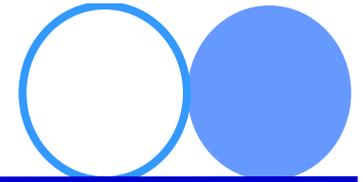
- 今までの因子分析は「探索的因子分析」
 - 観測変数を規定する因子についての仮説を設定せず，因子を探索的に検討するための因子分析
- 確証的因子分析
 - 因子について何らかの仮説があり，それが実際のデータにあてはめることができるかどうかを検討する因子分析
 - ある尺度の因子がすでに探索的因子分析で抽出されていて，尺度のデータを改めて得たときに，因子の仮説が妥当と言えるかを再度検証する
 - 確認的因子分析，検証的因子分析とも



各因子がどの観測変数に大きく影響するかをあらかじめ仮説として立てておき，分析する



因子間の関連を想定するので，必然的に「斜交回転」の因子分析になる



- どう分析する？

因子の仮説は、「因子(潜在変数)→観測変数」という因果モデルとして考えることができる

- 仮説的因果モデルがどれほど実際のデータを説明するか、分析できる方法が必要

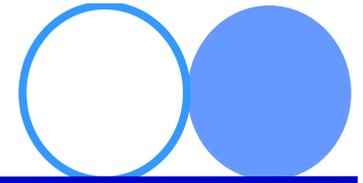
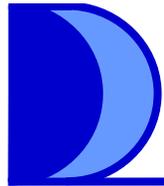
- 共分散構造分析(構造方程式モデリング:SEM)

- 因果モデルを検証するための多変量解析法

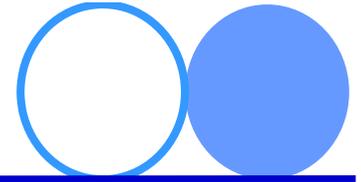
- 「潜在変数→観測変数」(測定方程式)
- 「観測変数or潜在変数→潜在変数」, 「観測変数→観測変数」(構造方程式)

→ 観測変数・潜在変数間の因果関係を検討

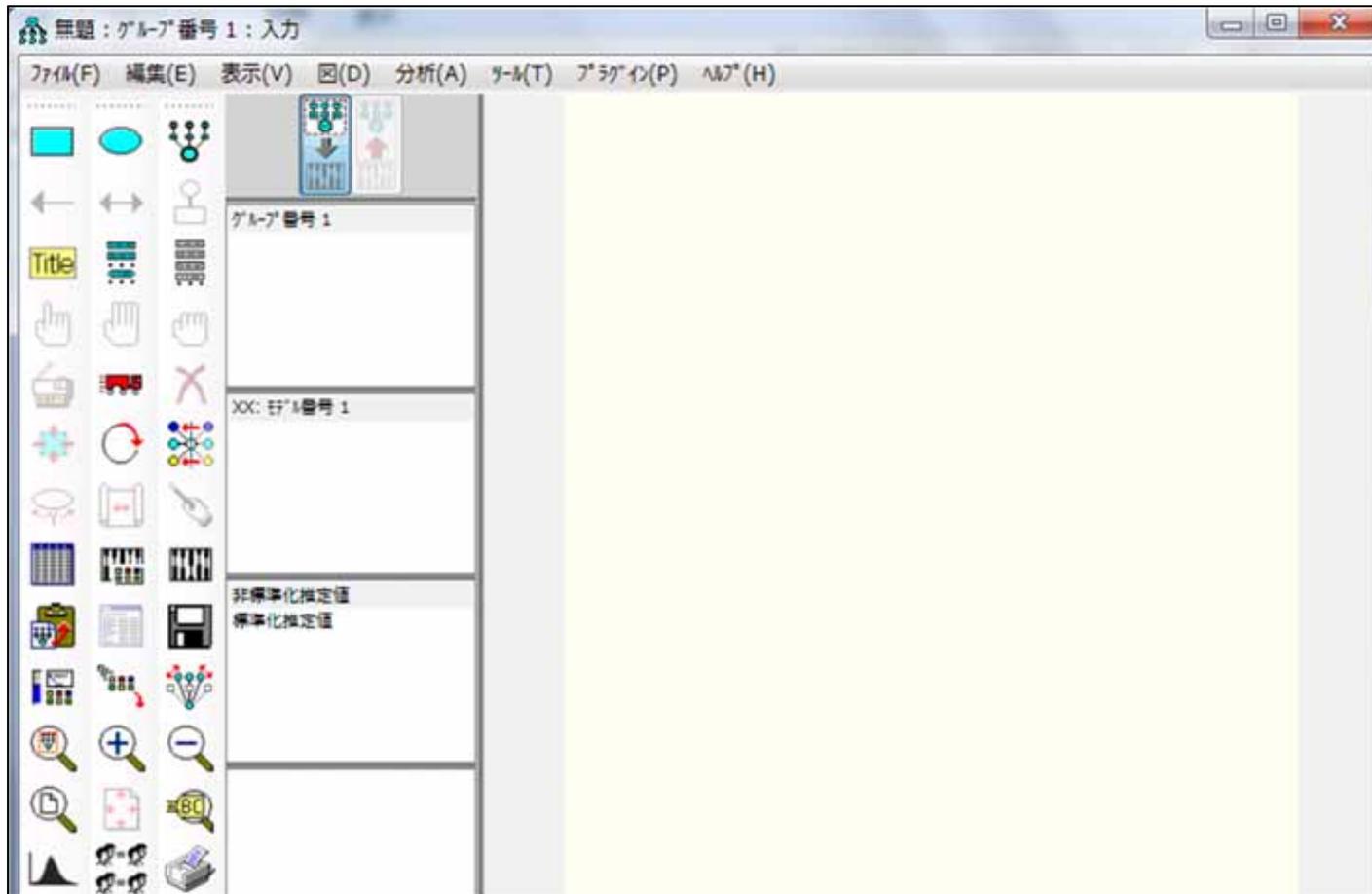
- 測定方程式を使えば、因子分析が可能

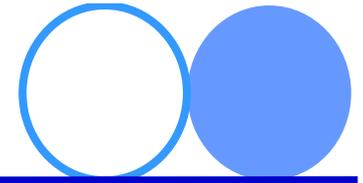


- Amos
SEMを行うための専用ソフトウェア
 - グラフィカル・インターフェイスを採用
 - 四角, 楕円, 矢印を使ってパス図を描画するだけで, 因果モデルを構築することができる
 - 構築した因果モデルについて, その係数やデータへの適合度(当てはまりの良さ)などを自動で計算してくれる
- SPSSのデータを直接読みこみ, 分析することができる(もちろんExcelからも可能)



- Amosを起動





- 因子分析モデルを描く

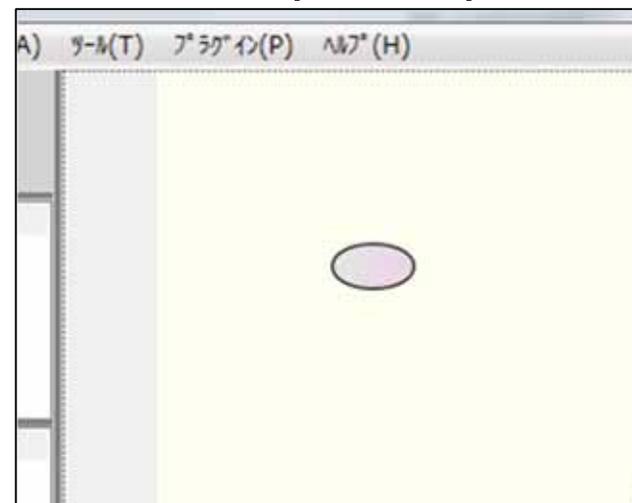
- ウィンドウ左側に並んだアイコンのうち, 一番右上のアイコン()をクリック

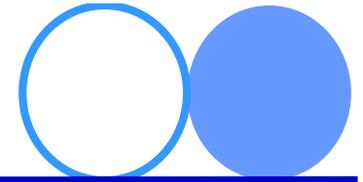
- 「潜在変数を描く」と表示されるアイコン

- カーソルがアイコンの形になる

- カーソルを右側の黄色いフィールド(描画領域)に合わせ, 適当な大きさの楕円を描く(下図)

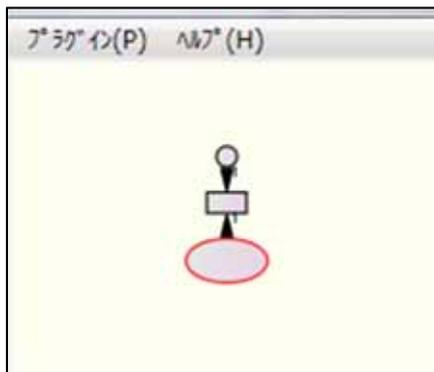
- 最終的にできるモデルが, 描画領域をはみ出さないように注意



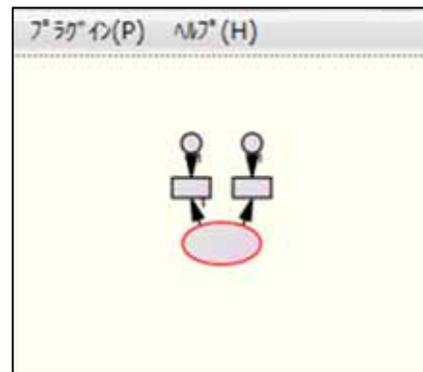


- 描いた楕円の上にカーソルを合わせたままクリックすると、変数が追加される

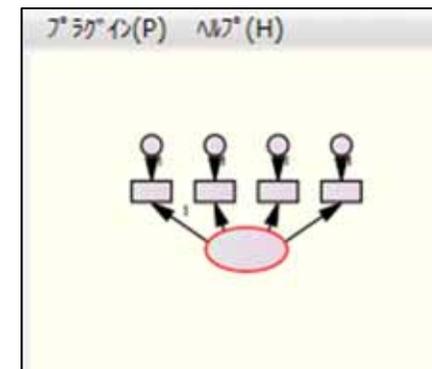
1回クリック

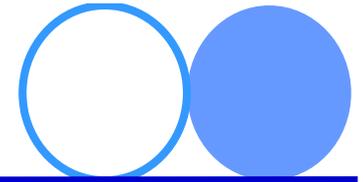


もう1回クリック

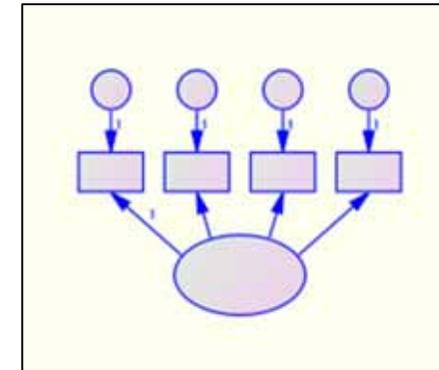


- 変数が4つになるまで続ける

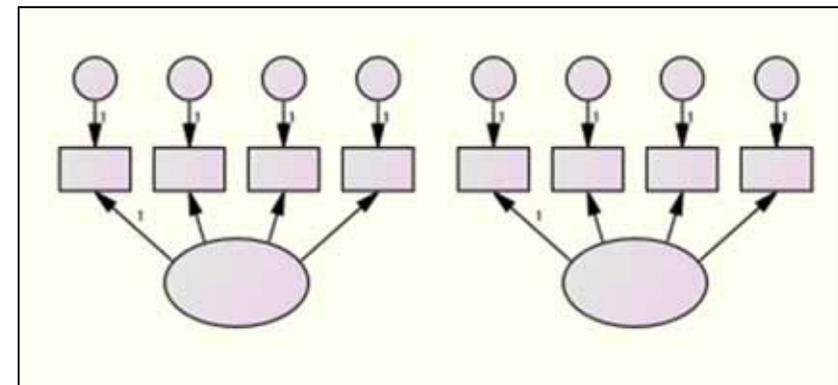


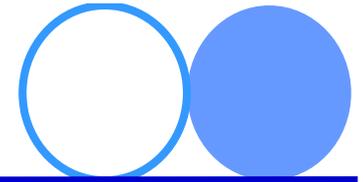


- できたパス図をもう一つ作る
 - 左側のアイコンから「」をクリック
 - パス図が全部青になる(右図)

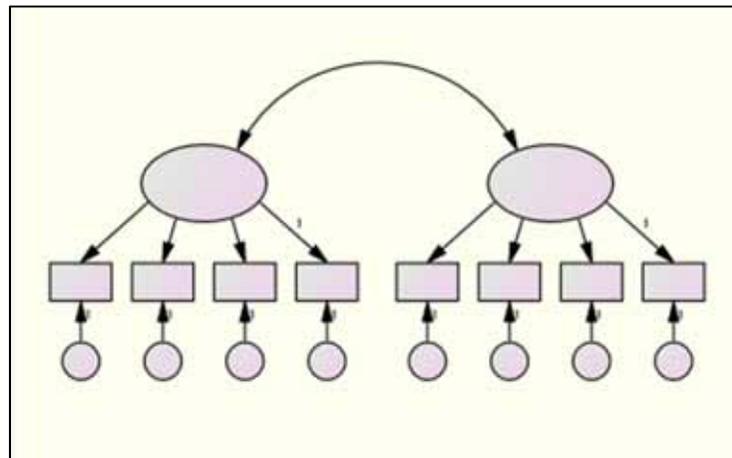


- アイコン列から「」を選択し,パス図上でクリックし,そのまま空白のところまでドラッグ
 - 同じパス図がコピーされる
- アイコン「」をクリック
 - 青が解除される(右図)

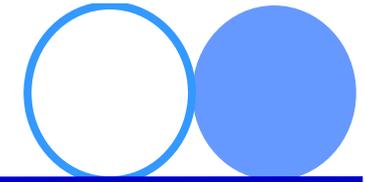




- パス図の向きを変える
 - アイコン列から「」を選択
 - 楕円上でクリックするたび、クリックした図を中心に右に90度回転 → 両方変えておく
- 共変関係の矢印を追加する
 - アイコン列から「」を選択
 - 楕円どうしを双方向矢印でつなぐ(下図)



パス図にデータを代入(1)



- データの読み込み

- 「」をクリック

- 「ファイル名」からデータファイルを選択
- Excelファイルの場合、ここでシートを指定する
- 「グループ番号1」「SHEET1(XLS)」と表示されたのを確認して「OK」

- 因子にデータ名を入力

- 左側の楕円をダブルクリック

- 変数名に「活動性」と入力し閉じる

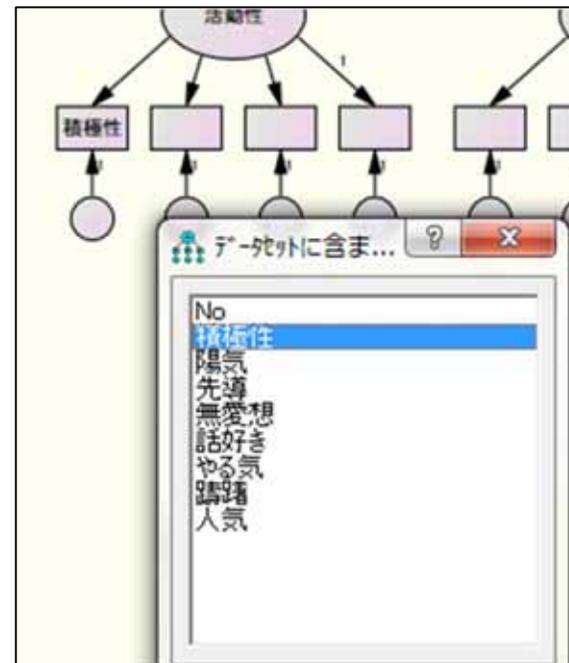
- もう一方の楕円には「社交性」と入力

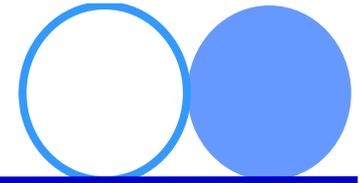


- 観測変数にデータを代入

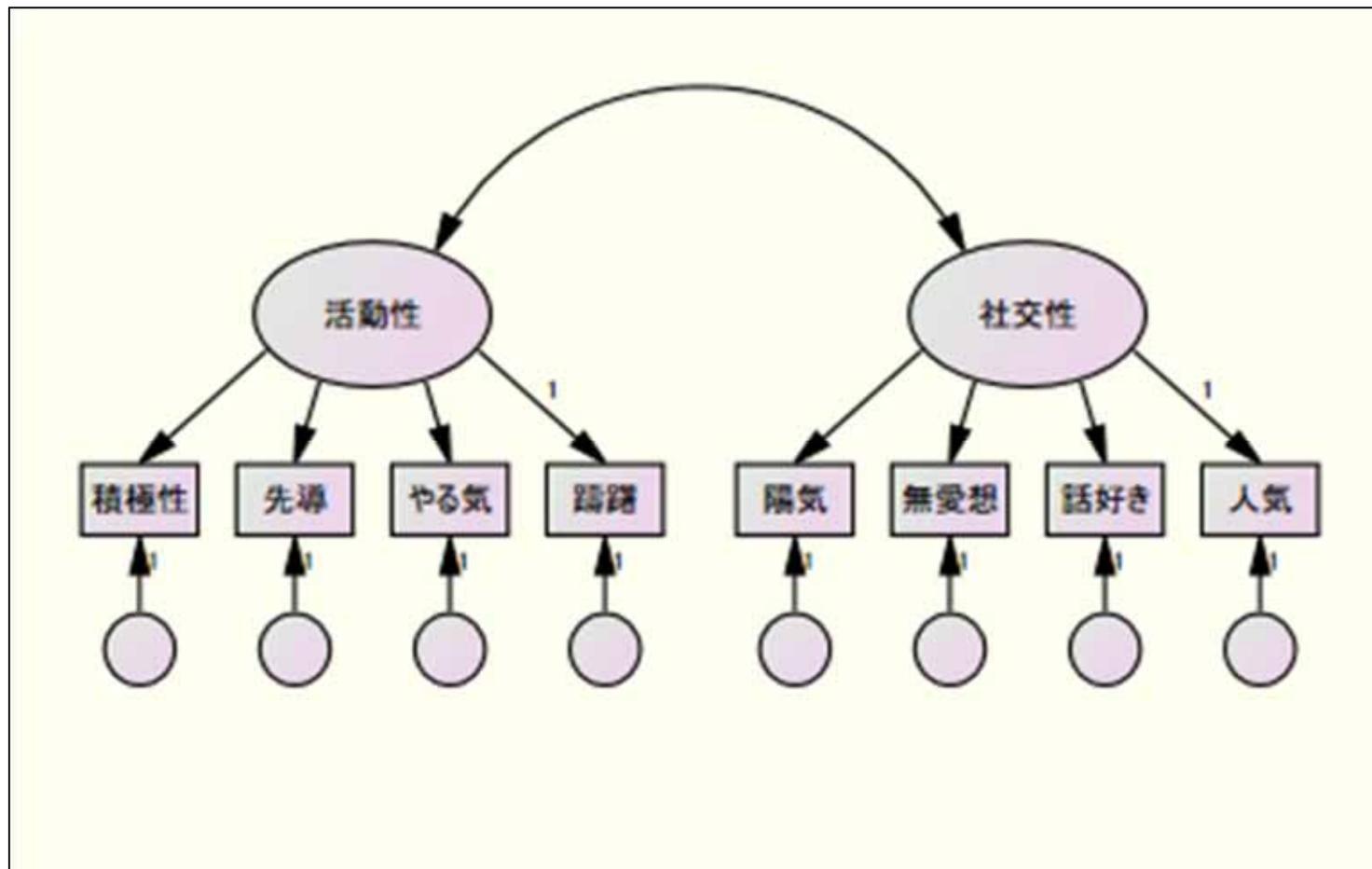
- 「」をクリックし、表示されるダイアログボックス(右下図)から、「積極性」を選択
- 一番左の観測変数までドラッグ
 - 変数内に「積極性」と表示

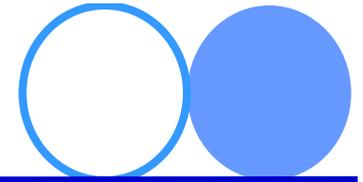
- 同様の操作を続ける
(見本は次スライド)



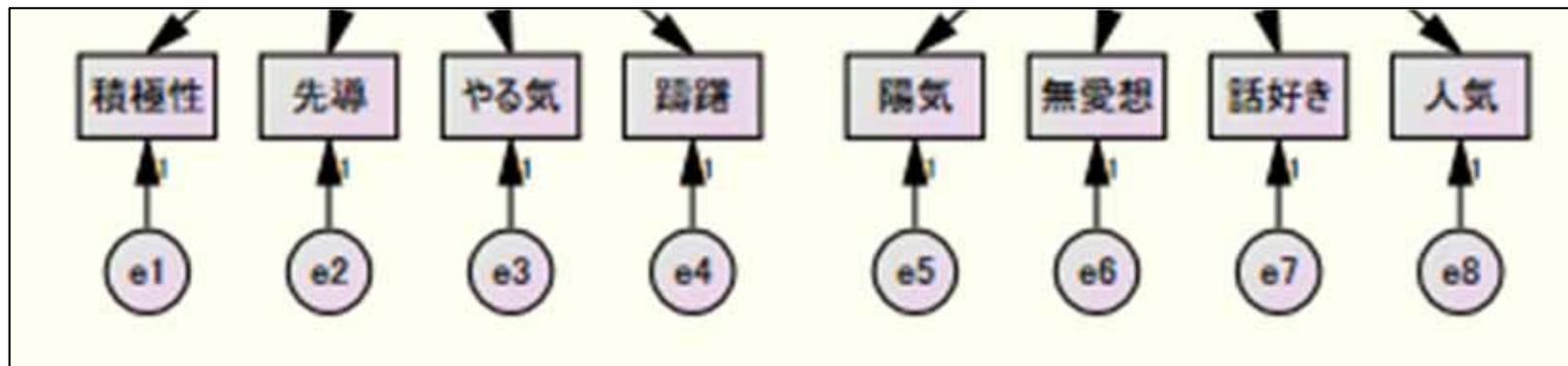


- 因子と観測変数の関係

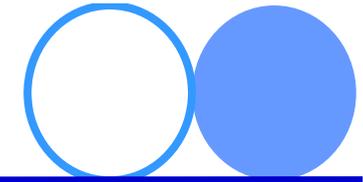




- パス図下段の円は「**誤差変数**」
 - 因子分析で「**独自性**」と言っていたのと同じもの
 - 因子名を入力したときと同じように、変数名を入力
 - 左から順番に「e1」「e2」・・・としておく(下図)



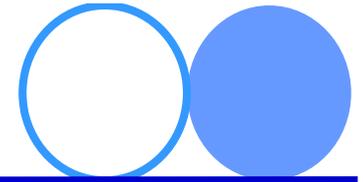
- **パス図の完成！**



- 分析の設定
 - 分析のプロパティ () をクリック
 - 「出力」タブで、「最小化履歴」と「標準化推計値」にチェックして閉じる

- 分析
 - 推計値の計算 () をクリック
 - 保存ダイアログが出るのでパス図を保存しておく
 - 中央下ウィンドウに「最小値に達しました」と出たら分析終了





出力パス図の表示

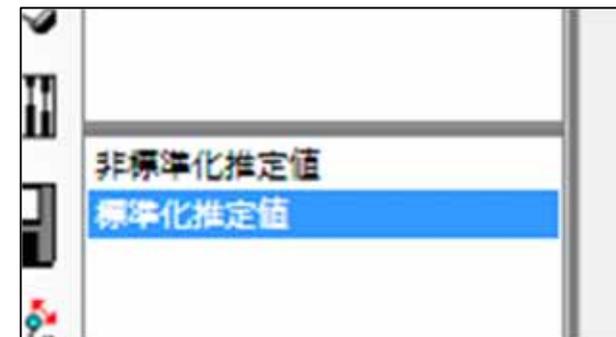
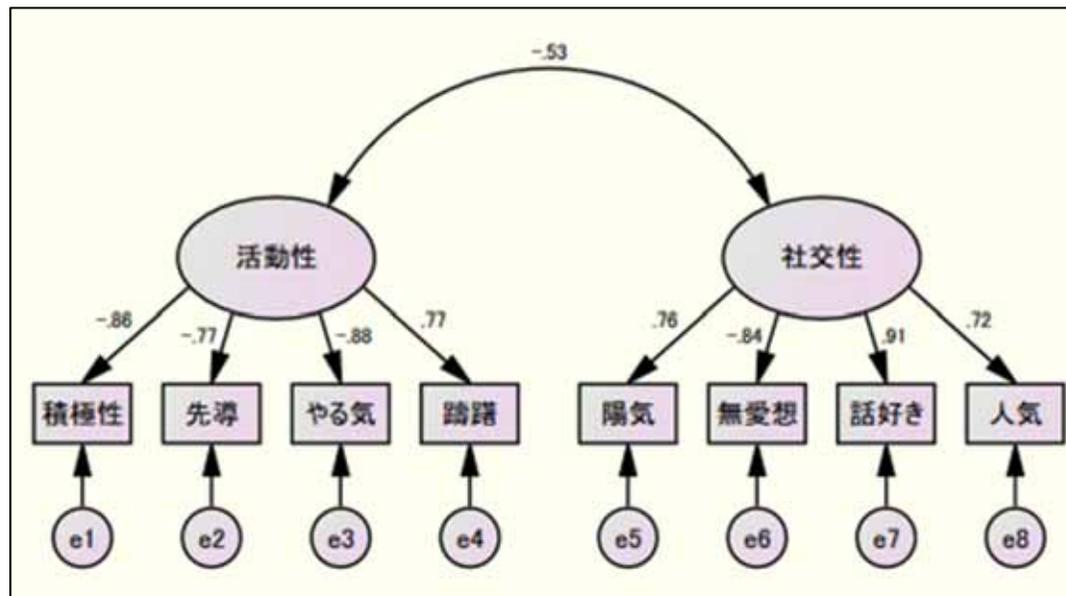
– 中央上のアイコンの右側をクリック

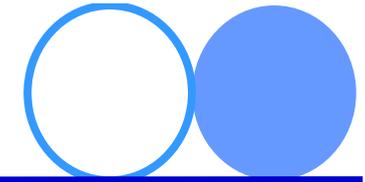
- パス図に数値が追加される



– 中央の「標準化推定値」を選択

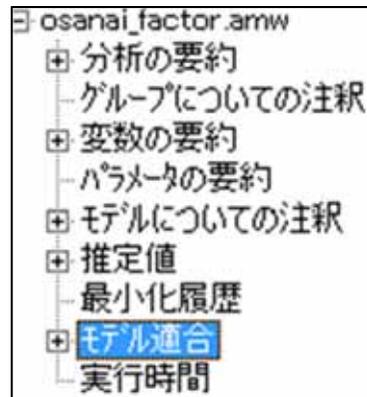
- 各因子から観測変数へのパスの数値 = **因子負荷量**
- 各因子の相互パスの数値 = **因子間相関**





• 適合度をみる

- 「」をクリックして、テキスト出力を表示する
- ウィンドウ左側の「モデル適合」を選択



モデル適合の要約

CMIN

モデル	NPAR	CMIN	自由度	確率	CMIN/DF
モデル番号 1	17	21.403	19	.315	1.126
飽和モデル	36	.000	0		
独立モデル	8	489.587	28	.000	17.485

RMR, GFI

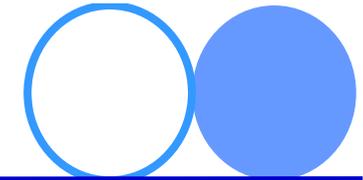
モデル	RMR	GFI	AGFI	PGFI
モデル番号 1	.120	.953	.911	.503
飽和モデル	.000	1.000		
独立モデル	1.009	.353	.169	.275

基準比較

モデル	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
モデル番号 1	.956	.936	.995	.992	.995
飽和モデル	1.000		1.000		1.000
独立モデル	.000	.000	.000	.000	.000

その上の「推定値」では、
パス図に表示された数値の
一覧表がみられる

- 数値は「モデル番号1」をみる



– CMINの確認

- 「確率」が.05以上なら、「構成されたパス図は正しい」という帰無仮説は棄却されない

モデル	NPAR	CMIN	自由度	確率	CMIN/DF
モデル番号 1	17	21.403	19	.315	1.126
飽和モデル	36	.000	0		
独立モデル	8	489.587	28	.000	17.485

– GFI, AGFI, CFIの確認

- 適合度指標。いずれの数値も0.9以上あれば、通常はモデルへの適合は良いと判断される

モデル	RMR	GFI	AGFI	PGFI
モデル番号 1	.120	.953	.911	.503
飽和モデル	.000	1.000		
独立モデル	1.009	.353	.169	.275

– RMSEAの確認

- 0.05未満であれば、モデルへの適合がよいと判断される

モデル	RMSEA	LO 90	HI 90	PCLOSE
モデル番号 1	.036	.000	.098	.583
独立モデル	.408	.377	.440	.000

文献(初回に紹介のなかった本を中心に)

- 足立浩平 (2006). 多変量データ解析法—心理・教育・社会系のための入門 ナカニシヤ出版
 - 多変量解析についての解説書。因果モデルを中心にわかりやすい構成。データを使った分析例も。
- 柳井晴夫・岩坪秀一 (1976). 複雑さに挑む科学—多変量解析入門 講談社
 - 重回帰分析から因子分析, 数量化理論まで, 多変量解析法全般について解説。出版年は古いが, 安価な割に説明はかなり詳しい。縦書き。
- 豊田秀樹・前田忠彦・柳井晴夫 (1992). 原因をさぐる統計学—共分散構造分析入門 講談社
 - SEMのしくみを詳しく解説。豊富な分析例を交えており, やかりやすい。安価。
- 豊田秀樹(編著) (2007). 共分散構造分析[Amos編] 東京図書
 - Amosについて詳しく解説。わかりやすい。
- 小塩真司 (2004). SPSSとAmosによる心理・調査データ解析—因子分析・共分散構造分析まで 東京図書
 - 心理学に関連する統計手法の多くを網羅。SPSSとAmosを使っでの分析法を紹介。わかりやすい。
- 緒賀郷志 (2010). Rによる心理・調査データ解析 東京図書
 - 上記の姉妹本。Rについてのわかりやすい入門書。ほとんどの統計手法を扱う。
- 青木繁伸 (2009). Rによる統計解析 オーム社
 - Rについての詳しいテキスト。かゆい所に手が届く。ただ上級者向け。著者のサイトにはRの関数が大量にアップされている。